DOCUMENT RESUME

ED 172 217 **CS 204 869**

AUTHOR Freedman, Sarah Warshauer
TITLE A Comparative Evaluation of Student and Professional Prose.
PUB DATE Apr 79
NOTE 16p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, California, April 8-12, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adults; *Authors; *College Students; Comparative Analysis; *Composition (Literary); Educational Research; *Evaluation Criteria; Higher Education; Language Usage; *Literary Styles; *Student Evaluation

ABSTRACT

The hypothesis that professional writers would be judged superior to college students on both holistic and analytic evaluations was only partially confirmed when four teacher-evaluators rated the anonymously submitted compositions of 64 college students and 5 professional writers. On the holistic scale the professionals did not distinguish themselves, although they were rated more highly than the average student (2.65 versus 2.25 on a four-point scale). The professionals did score significantly higher on the analytic scale than the students. Whereas the students received the three highest scores on the holistic scale, the professionals received the three highest scores on the analytic scale. Similarly, the college students usually received the same score regardless of the scale used, while the professionals received consistently higher analytic ratings. The main difference between the two groups' writing styles seemed to be "register," in that students did school writing for a grade by the teacher, and the professionals were not concerned about grades and wrote in a world writing register, aiming at a diverse audience. For this reason, the results suggest that the register of school writing needs to be examined critically and modified by the teaching profession to make it less distinct from world writing, which is the true goal of communication. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED172217

A Comparative Evaluation of Student and Professional Prose

Sarah Warshauer Freedman
English Department
San Francisco State University
San Francisco, CA 94132

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Sarah Warshauer Freedman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented at American Educational Research Association
San Francisco
April, 1979

CS204869

A Comparative Evaluation of Student and Professional Prose[1]

A laudable goal espoused by many teachers of writing is for their students to learn to approximate the style and proficiency of the professional writer. To achieve this goal, writing teachers frequently expose their students to professional models, having them read exemplary essays and sometimes asking students to imitate these essays directly. Writing teachers may also create syntax lessons to teach structures typical in professional prose. Such lessons frequently focus on the cumulative sentence discovered by Francis Christensen (1967).

On the other hand, disquieting research has revealed that many of the techniques for teaching writing, advanced in popular rhetoric texts, have little relation to the practices of published professional writers. Meade and Ellis (1970) showed that most traditional rhetorical methods of paragraph organization (e.g., description, cause-effect, comparison, definition) were not usual methods of paragraph organization for the professional sample they studied. Braddock (1974) showed that the topic sentence, the stock opening line for the "adequate" student paragraph, was difficult if not impossible to locate in a braod sample of paragraphs by professional writers.

In this study I shall look for the differences between the professional and the student in hopes of advancing the goal of teaching students to write as professionals do. Instead of examining the differences between published professional prose and the dicta of the rhetoric text, as Meade and Ellis and Braddock did, I shall compare professional prose with student prose. But, instead of comparing only those samples of professional prose which have been

published with student prose, as Christensen did, I shall compare the performance of the student and the professional on identical expository writing tasks performed in almost identical writing situations. Whereas Christensen sought in part to define how the student is deficient in meeting a final goal, published professional prose, I shall use the criterion of impromptu professional writing to examine what one might expect of the student writer on an impromptu writing task.

I designed my study so that I would be able to measure the differences between the two groups in two ways:

(1) with a comparison of both holistic and analytic ratings assigned by expert teachers to the papers of both professionals and students,

(2) with a qualitative analysis of the differences between the student and professional samples, designed to explain the meaning of the ratings.

Past researchers have not compared teacher ratings of student and professional prose. By including the teacher's judgments, I hoped to glean information about the basis of the teachers' evaluations and about teachers' expectations for good student prose. The teachers thought they were judging only student writing. I proposed the following hypothesis: teacher-judges, who would not know they were rating professional prose as well as student prose, would award consistently higher ratings to professional prose on both holistic and analytic rating scales. If any part of the hypothesis was not proved, I hoped that any consistent differences between the essays of the professionals and those of the students found upon critical analysis of the samples would provide grounds for speculation about either the bases of the teachers' ratings or about the abilities of the professional writers given this particular writing task.

Eight professional writers from various parts of the United States were recruited to volunteer to participate in the study. Although all eight agreed to participate, only five actually wrote essays for the study. I agreed to keep the identity of the professional writers confidential, but the following information indicates their caliber. All have published books and articles; all have had extensive experience teaching composition at the college level. One is a novelist and poet; one is a literary scholar and the author of a freshman rhetoric text and has directed a freshman composition program at a major university; one is an eminent researcher on composition; another has authored a best selling text on teaching writing and is a literary scholar.

Sixty-four student writers were selected from required writing classes at four San Francisco area colleges ranging in type from highly select, private schools to open-admissions, public schools. Thus, the student writers exhibited a wide range of abilities. All were volunteers.

Four teachers on the staff of Stanford University's freshman writing program were selected to be judges of the essays in this study. The judges all had had at least three years experience teaching college writing, and all were recommended for their expertise in essay evaluation by the director of the Stanford program. All were paid.

Topics asked the writer in 45 minutes to compare and contrast two quotations or to argue his or her opinion on a controversial subject. In all, there were eight topics, four of each type. A sample from each type of topic follows:

1. A Founding Father said: "Get what you can, and what you get hold; 'Tis the Stone that will turn all your Lead into Gold.'"

A contemporary writer said: "If it feels good, do it."

What do these two statements say? Explain how they are alike and how they are different.[2]

2. President Ford gave Nixon an "unconditional pardon." Do you agree or disagree with Ford's decision? Give reasons for taking your position.

The classroom teachers in two different classes at each of the four schools, eight classes with eight different teachers in all, distributed the eight essay topics randomly. One essay on each topic from each class was selected for the study, making sets of eight student essays on each topic, a total of 64 essays by the student writers.

Unlike the students, the five professionals did not write during a class. Each wrote on a different topic which was mailed to him or her by the investigator. All received instructions not to look at the topic until they were ready to spend 45 minutes writing. They were further instructed to write their essays during one sitting, as though they were in class. Professionals were allowed to compose on a typewriter; three of the five did.

After the 64 student essays and the five professional essays were collected there were eight student essays on every topic and one additional professional essay on five of the topics. After collecting the essays, I coded them with numbers and typed them to conceal the identities of the writers. Furthermore, typing removed the potential effect handwriting might have on the judges (Markham, 1976). The essays were transcribed exactly, including all errors.

The four judges rated the essays on two Saturdays, two weeks apart. All four gave holistic ratings to each essay. Two of the four gave analytic ratings to each essay, a single rater giving an analytic rating to only half of the sample. Before the evaluation sessions began, the judges received a written paragraph about the research which stated that judges would be evaluating 45 minute essays on assigned topics written mostly by Bay Area college

students.

The judges rated holistically first, rating essays on each of the eight topics separately. Before giving their holistic ratings for each topic, the judges received the usual holistic training, practice in applying the holistic rating scale to sample essays. This training procedure was repeated before ratings on every topic, with the sample essays being on the topic to be read next.

After discussing the sample training essays, the judges received their reading packet containing the experimental student and professional essays for the topic. In these reading packets several optional training essays preceded the experimental essays. The judges thought that all essays in the reading packet were the experimental student essays, but the optional training essays were placed in the packet in case any of the judges evidenced the need for additional training. They also were used to insure that the reading packets on every topic always consisted of the same number of essays. If there was no professional essay on a topic, three optional training essays preceded the eight experimental student essays; if there was a professional essay in addition to the student essays, only two optional training essays preceded the nine experimental essays. So every reading packet contained eleven essays, and the raters thought all eleven were produced by the experimental student group. In the group of actual experimental essays, the eight student essays always came first, in a different random order for each judge, and the professional essay always came last to keep judges from elevating their expectations for the students' essays.

After the raters had given their holistic ratings on all eight topics, they were asked two weeks later to provide analytic ratings for half of the

essays they had already rated holistically, essays on four of the topics. Before the analytic rating on each topic, judges were retrained. The analytic training differed from the holistic training only in that judges practiced applying the analytic scale rather than the holistic scale.[3] The same training essays and optional training essays were used again. To use the analytic scale, judges had to give reliable ratings in six different categories: voice, development, organization, sentence structure, word choice, and usage

To summarize, all four judges rated all essays on all eight topics with the holistic scale. However, only two judges rated the papers on each topic with the analytic scale. So each paper received four holistic and two analytic ratings.

The hypothesis, that professional writers would be judged as more outstanding than students on both rating scales, was only partially confirmed. On the holistic scale the professionals did not distinguish themselves as outstanding although they were rated more highly than the average student. The average score given a professional on the four-point holistic scale was 2.65; the average score given a student was 2.24. Four was the best score and one the worst. According to a t-test, the differences in the mean holistic scores for the professional versus the student groups were insignificant ($t = 1.19$). Although professional holistic scores were slightly better, on the average, students received the three highest holistic scores.

On the analytic scale the professionals fared better. This scale contained six categories with a possible score of six for each category. Thus, if one sums the scores across all of the categories, a writer could receive a maximum score of 36 from each of the two judges. The average summed analytic score given the professionals was 30.2; the average for the students was 19.5. According to a t-test, the differences in the mean analytic scores for the

professional versus the student groups were significant at the .02 level of confidence (t = 2.43). Whereas on the holistic scale students received the three highest scores, on the analytic scale professionals received the three highest scores.

A breakdown of the analytic scale by categories reveals that the professionals proved outstanding on voice, sentence structure, word choice, and usage. But they were not judged so consistently high on the categories development and organization. The boost in the analytic score came mostly from high scores on the more technical, style oriented categories (sentence structure, word choice, and usage) and from the style/personality category (voice).

A final interesting difference between the scores given the professionals versus those given the students was that students usually received the same score regardless of the rating scale while professionals did not. The professional papers would receive consistently higher scores when rated analytically than when rated holistically. Thus, those professionals who received the same ratings as the students on the holistic scale could expect to receive ratings higher than their student counterparts on the analytic scale. Table 1 compares the analytic scores for each professional paper with the range of analytic scores for the student papers that received holistic scores identical with the professional's. In Table 1 the scores are summed across all readers.

Insert Table 1 about here

So, for the holistic rating, across four raters, with the top score given by each rater being a four, a paper could receive a high holistic score of 16, if it received a four from all four raters. And on the analytic scale,

a paper could receive as many as 36 points from each of two raters for a maximum of 72 points. Notice, in Table 1, that the professional papers that received a summed holistic score of nine received summed analytic scores of 55 and 59; the student papers receiving a holistic score of nine received analytic scores between 34 and 54. Actually, only one of these students received a 54; the others received socres below 49. The other holistic/analytic comparisons in Table 1 bear out the discrepancy between students' and professionals' holistic and analytic scores.

The reason for the professional writers' unexpectedly low scores on the holistic scale and on the development and organization categories of the analytic scale deserves comment. After a critical analysis of the content categories of the professional essays, I found them generally superior to the students'. I agreed with the results of the summed analytic rating; the professionals seemed to distinguish themselves. My analysis of the professionals' papers revealed, however, that they consistently followed a different set of appropriateness conditions than the students did, approaching the writing task differently in at least three ways. I propose that these differences in approach influenced the judges to penalize them on the development and organization categories of the analytic scale and also to lower their holistic scores.

First, the professionals wrote on these formal topics personally, with less distance than the student. What student would dare begin an essay, "First I want to answer 'damned if I know.' Then 'who cares?'" Unless the student was signalling to the teacher that he or she was resigned to a failing grade. But the professional writer who began this way went on to deal with the topic in a truly insightful and creative way:

> What is more interesting to me than the answer is the reasoning I'm forced to go through to achieve my considered indifference. For the issue is full of things to which I'm not indifferent.

Second, three of the professional writers implied in their essays that they were alienated by their topics. They seemed not to be able to write without first overcoming this alienation. They spent time musing on paper as they tried to find something within the topic that was personally meaningful, an honest approach.

Third, one of the professionals liked his topic but used many scholarly references to support his opinions--Forester, Schumacher, Mills, and Rawls. One judge reported penalizing this writer for "obnoxious name dropping."

The professional essays seemed to upset the judges' expectations and to cause controversy. Their style was superior, so judges took off for their content. The stylistic categories showed the professionals superior on the analytic scale, but those content categories of development and organization which caused judges to lower their analytic scores probably also caused them to lower their holistic evaluations.

These differences between professional and student writers seem equivalent to differences between registers in speech. I shall label the register for student writing <u>school writing</u>. School writing usually consists of prose written for a teacher audience with the idea that the teacher will evaluate the prose. The usual reason given the student for performing the writing task is to receive instruction. This writing situation necessarily places the student writer in a subordinate role to the evaluator who possesses the power. Rarely is the audience wider than the teacher, and then it

most commonly includes only other members of a writing class. I shall label the register for professional writing world writing. World writing usually consists of prose written for an audience of strangers. The audience for school writing is never as diverse as the one the professional writer writes for. The professional writes to communicate ideas. Although she or he expects the audience to evaluate the ideas, the professional generally has more authority and power than the audience. The professional, not being subordinate to the audience, writes with an air of authority. Because the professionals in this study did not switch from world writing to school writing, the judges penalized them.

It seems to me that teachers subconsciously and unintentionally, like the judges in this study, do many students a disservice by penalizing them for doing world writing. Indeed, once the students' audiences stretch beyond the bounds of the classroom, the register should no longer be school writing. Just as the professional writer in this study did not switch easily into the subservient role of the author in school writing, I hypothesize that the students will not be able to switch so easily into the confident role of the professional writer who aims to instruct or inform rather than to demonstrate knowledge for the purpose of receiving instruction or of pleasing the teacher. And just as the professional writer was penalized for assuming an inappropriate role when the evaluator expected school writing, so the student will never be accepted as a professional writer without learning to assume the role of the professional writer. Certainly, teachers of college level and other advanced writers need to think about how to help their students make the transition from a student to a peer, and furthermore, teachers must develop a new set of appropriateness conditions for both the writer and the judge of the writing.

Most writing teachers teach students to write because they believe students need to be able to write outside the confines of the classroom. Writing teachers must begin to pay attention to the differences between student and professional writing so that they can help students make the transition between student writing and world writing. Or even more radically, teachers need to allow students to take more authority when they write. The register of school writing needs to be examined critically, and perhaps consciously modified by the teaching profession so that school writing will be less distinct from world writing which has a true goal of communication.

References

Braddock, R. The frequency and placement of topic sentences in expository prose. Research in the Teaching of English, 1974, 8, 287-302.

Christensen, F. A generative rhetoric of the sentence. Notes Toward a New Rhetoric. New York: Harper and Row, 1967, 1-22.

Markham, L. Influences of handwriting quality on teacher evaluation of written work. American Educational Research Journal, 1976, 13, 277-283.

Meade, R. and Ellis, W.G. Paragraph development in the modern age of rhetoric. English Journal, 1970, 59, 219-226.

Footnotes

[1]This article is based on part of my doctoral dissertation, Influences on the Evaluators of Student Writing, Stanford University, 1977. Special thanks go to Professor Robert Calfee who guided my research and to Professor Elizabeth Traugott for her invaluable help on the section of the study leading to this paper.

[2]This topic was first developed by the California State University and College System for use in their Freshman English Equivalency Test.

[3]The analytic scale, developed for this study, is an adaptation of the Diederich and Adler scales and is available upon request.
(Diederich, P. Measuring growth in English. Urbana, Ill.: National Council of Teachers of English, 1974.
Adler, R. An investigation of the factors which affect the quality of essays by advanced placement students, Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, 1971.)

TABLE 1

PROFESSIONAL VERSUS STUDENT SCORES

Holistic Score	Analytic Score Range	
	Professional	Student
4	---	22 - 28
5	---	12 - 38
6	---	31 - 42
7	---	26 - 42
8	---	24 - 47
9	55 - 57	34 - 54
10	59	27 - 55
11	67	37 - 57
12	---	39 - 49
13	---	43 - 58
14	64	52
15	---	47 - 56